



**RUI AMÉRICO
FERREIRA DA COSTA**

**PROPRIEDADES DA CENTRALIDADE EM REDES
COMPLEXAS: DISTRIBUIÇÃO E DEPENDÊNCIA COM
A CONECTIVIDADE**



**RUI AMÉRICO
FERREIRA DA COSTA**

**PROPRIEDADES DA CENTRALIDADE EM REDES
COMPLEXAS: DISTRIBUIÇÃO E DEPENDÊNCIA COM
A CONECTIVIDADE**

dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Física, realizada sob a orientação científica do Dr. José Fernando Mendes, Professor catedrático do Departamento de Física da Universidade de Aveiro

Dedico este trabalho aos que, durante a infância, lançaram a semente da sensibilidade e curiosidade no meu espírito, e aos que ao longo dos anos a incentivaram, e incentivam, a crescer e frutificar.

o júri

presidente

Prof. Dr. José Fernando Ferreira Mendes

professor catedrático do Departamento de Física da Universidade de Aveiro

Prof. Dr. Fernão Rodrigues Vístulo de Abreu

professor auxiliar do Departamento de Física da Universidade do Aveiro

Prof. Dra. Maria do Céu Almeida de Moraes Marques

professor associado do Departamento de Física da Faculdade de Ciências da Universidade do Porto

palavras-chave

grafo, rede *scale-free*, lei de potência, árvore, caminho mais curto, conectividade, centralidade de intermediação, .

resumo

Em redes *scale-free* a conectividade exibe uma distribuição em lei de potência com expoente $\gamma \in]2, \infty[$. Em contraste com o carácter local desse parâmetro a distribuição de BC (*Betweenness Centrality*) é definida globalmente e parece ser universal em cada uma de duas classes de redes [12,13]. As particularidades dos grafos tipo árvore permitem simplificar a definição de BC, e usando a aproximação ao contínuo deduz-se a relação entre conectividade e BC em árvores evolutivas *scale-free*. Mostrou-se que o expoente na relação $g \sim k^\eta$ [12,15] é para estes grafos $\eta = \gamma - 1$.

Este resultado simples, todavia intrigante, conduz quase imediatamente à distribuição de BC, que se conclui inequivocamente ser da forma $P(g) \sim g^{-2}$, independentemente dos detalhes da estrutura.

Estes resultados são testados computacionalmente para vários tamanhos de rede e expoentes γ e η , e ambos são corroborados pelas simulações.

Mostrou-se ainda analiticamente a relação entre a BC das arestas, referida em [13], e a BC convencional, isto é BC dos nodos.

keywords

Graph, scale-free network, power-law, tree, shortest-path, connectivity, betweenness centrality.

abstract

In scale-free networks connectivity exhibits a power law distribution with exponent $\gamma \in]2, \infty[$. In contrast with the local nature of this parameter, the globally defined BC (Betweenness Centrality) appears to be universal within each of two classes of networks [12,13]. The particularities of tree-like graphs simplify the BC definition, and by applying the continuum approach one derives the relation between connectivity and BC in *scale-free* evolving trees. One shows that the exponent of the relation $g \sim k^\eta$ [12,15] is in these graphs $\eta = \gamma - 1$.

This simple, nevertheless intriguing, result leads almost immediately to the BC distribution, which is clearly concluded to be in the form $P(g) \sim g^{-2}$, regardless of the structure details.

This results are computationally tested for a number of network sizes and exponents γ and η , and both are corroborated by the simulations. One still analytically derives the relation between the edge BC, referred in [13], and the conventional BC, that is node BC.

ÍNDICE

1 – INTRODUÇÃO	1
.1 - REDES EVOLUTIVAS	1
1.1.1 - ESTRUTURA <i>SCALE-FREE</i> (MODELO B-A)	3
1.1.2 - APROXIMAÇÃO AO CONTÍNUO	6
1.1.3 - CAMINHOS E DISTÂNCIA MAIS CURTA ENTRE VERTICES	10
1.1.4 – ÁRVORES	11
1.2 - MEDIÇÃO DA CENTRALIDADE	13
1.2.1 - CENTRALIDADE DE INTERMEDIÇÃO (BC)	15
1.2.2 - FUNÇÃO DISTRIBUIÇÃO DE BC	16
1.2.3 - DEPENDÊNCIA COM CONECTIVIDADE	17
1.3 - MODELO COMPUTACIONAL	18
1.4 – ALGORITMO PARA O CÁLCULO DA BC	19
2 - BC EM ÁRVORES <i>SCALE-FREE</i>	24
2.1 - TRABALHO ANTERIOR	25
2.2 - EQUAÇÕES DIFERENCIAIS	27
2.3 - CÁLCULO DA BC	29
2.4 - DISTRIBUIÇÃO DE BC	30
3 - SIMULAÇÕES E ESTATÍSTICA	33
4 – DISCUSSÃO	41
5 - CONCLUSÕES	44
6 - ANEXO	45
7 – REFERÊNCIAS	48

1-INTRODUÇÃO

1.1-REDES EVOLUTIVAS

Qualquer rede (grafo) é composta essencialmente por dois tipos de elementos, nodos e ligações (vértices e arestas), onde cada ligação une dois nodos tornando-os vizinhos. Pode pensar-se nisto como num número de pontos unidos entre si através de linhas, o número de ligações, ou conectividade, de um nodo é geralmente chamado o seu grau. No contexto da modelação os vértices desempenham o papel dos agentes do sistema, ao passo que as arestas representam as interacções entre agentes.

Uma rede evolutiva não é mais que uma rede que cresce, à medida que o tempo passa, com a chegada de novos nodos que criam novas ligações com os nodos previamente existentes de acordo com determinados princípios. Deste modo os nodos mais antigos concentram mais ligações que os mais recentes, o que produz heterogeneidade na distribuição das arestas no grafo [2]. Em contraste com os grafos estáticos, que tendem para estados de equilíbrio, estes sistemas não se encontram em equilíbrio.

De entre os principais exemplos de redes evolutivas destacam-se as populares Internet e a World Wide Web. Elas têm um tremendo impacto nas nossas vidas; a sua presença e utilização universalizada, os variadíssimos potenciais e a simplicidade de utilização contam-se entre as razões porque não poderíamos passar sem elas. Seria de esperar que se tivesse um conhecimento mais profundo de plataformas tão proeminentes e com tanto interesse (social, comercial, cultural, tecnológico, científico, etc). Mas, surpreendentemente, sabe-se muito pouco da sua estrutura e organização hierárquica, da topologia global, propriedades locais e de variados processos que ocorrem no seu interior. Este conhecimento é indispensável para otimizar o funcionamento da Internet e da WWW, para garantir segurança e explorar todas as suas possibilidades. Este tipo de problemas é objecto de estudo da Física Estatística.

Porém estas redes admiráveis são apenas exemplos de uma numerosa família de redes evolutivas. Na Ref. [1] são dados vários exemplos, como: redes de colaborações, relações públicas, citações em publicações científicas, redes de transportes, muitas redes

biológicas, etc. O facto de estas redes terem tamanho limitado impõe graves restrições à recolha de resultados experimentais úteis, devido à forte influência dos efeitos de tamanho finito e, não raramente, a insuficiência de dados estatísticos.

Recentemente foram descobertas características gerais na organização estrutural de numerosas redes evolutivas existentes na Natureza [1], tornou-se claro as estruturas complexas *scale-free* que exibem são consequência natural dos seus princípios de crescimento.

1.1.1-ESTRUTURA *SCALE-FREE* (MODELO B-A)

As redes complexas dividem-se em categorias distintas, com propriedades distintas, de acordo com a distribuição de grau que apresentam. Quando uma rede tem função distribuição de grau $\left(P(k) = \frac{\text{nº de nodos com } k \text{ ligações}}{\text{nº total de nodos na rede}} \right)$ com a forma de uma lei de potência, diz-se ser uma estrutura *scale-free*. Ou seja, numa rede *scale-free* a probabilidade de um vértice aleatoriamente escolhido ter grau k é proporcional a $k^{-\gamma}$:

$$P(k) \sim k^{-\gamma} \quad (1)$$

Foi dito anteriormente que um grande número de redes evolutivas conhecidas na Natureza exhibe comportamento *scale-free*, o que naturalmente levanta a questão de como se auto-organizam estas redes enquanto evoluem. “Quais os mecanismos responsáveis por esta organização?” Em resposta a esta pergunta apareceu o conceito de *ligação preferencial*.

Se numa rede evolutiva os novos vértices se ligarem aleatoriamente aos preexistentes (sem algum tipo de preferência) surge uma distribuição de grau exponencial ($P(k) \sim e^{-k}$). O modelo Barabási-Albert (B-A) baseia-se na ideia de que nas redes reais a ligação entre os novos nodos e os preexistentes é feita com algum tipo de preferência. Concretamente, quando um novo nodo chega à rede e escolhe a que nodos antigos se ligará, a probabilidade de um determinado nodo ser escolhido será proporcional ao seu grau nesse instante. Por exemplo, quando se faz uma nova referência numa página web, a probabilidade de se referir a um documento muito conhecido é maior do que a de se referir a um documento obscuro que quase ninguém conhece e a que ninguém se refere. Os vértices populares, isto é com elevado grau, são mais atractivos para as novas arestas que os de baixa conectividade; a popularidade é atractiva.

O modelo B-A é o exemplo mais simples de ligação preferencial. Neste modelo temos a seguinte dinâmica:

- i) Em cada passo de tempo um novo vértice é adicionado à rede;
- ii) O novo vértice escolhe um dos vértices antigos, com probabilidade proporcional ao seu grau, para criar uma aresta que una os dois.

Neste modelo a preferência é linear, isto é probabilidade de um vértice antigo ser escolhido é linearmente proporcional ao seu grau. No tempo t o número total de arestas é t e a soma dos graus de todos os vértices é o dobro, $2t$, assim a probabilidade de um vértice antigo receber a nova ligação é $k/2t$. Esta forma de preferência é apenas uma forma particular de preferência, mas encontra-se em várias redes reais [3].

Chama-se equação mestra deste modelo à seguinte expressão:

$$p(k, i, t+1) = \frac{k-1}{2t} p(k-1, i, t) + \left(1 - \frac{k}{2t}\right) p(k, i, t) \quad (2)$$

onde $p(k, i, t+1)$ representa a probabilidade de o nodo i ter grau igual a k no tempo $t+1$. Em cada termo da expressão (2) está prevista uma das duas possibilidades: o primeiro termo desta equação é a probabilidade do nodo i ter grau $k-1$ no tempo t multiplicada pela probabilidade de receber a ligação seguinte, $\frac{k-1}{2t}$, e o segundo termo representa a probabilidade de ter grau k no tempo t multiplicada pela probabilidade de não receber a ligação seguinte, $1 - \frac{k}{2t}$. A condição inicial desta equação é $p(k, i=0, 1, t=1) = \delta_{k,1}$ e a condição fronteira $p(k, t, t) = \delta_{k,1}$. A abordagem da equação mestra é muito eficiente para o estudo das redes em geral.

A distribuição de grau da rede inteira no instante t é dada por:

$$P(k, t) = \frac{1}{t+1} \sum_{s=0}^t p(k, i, s) \quad (3)$$

A aplicação de $\sum_{s=0}^t$ aos dois lados da equação (2) combinada com a definição (3) produz a equação mestra para a distribuição de grau total:

$$(t+1)P(k, t+1) - tP(k, t) = \frac{1}{2}[(k-1)P(k-1, t) - kP(k, t)] + \delta_{k,1} \quad (4)$$

No limite em que $t \rightarrow \infty$ (regime estacionário) a equação da distribuição de grau é:

$$P(k) + \frac{1}{2}[kP(k) - (k-1)P(k-1)] = \delta_{k,1} \quad (5)$$

No limite k contínuo a solução desta expressão toma a seguinte forma:

$$P(k) + \frac{1}{2} \frac{\partial[kP(k)]}{\partial k} = 0 \quad (6)$$

cuja solução é:

$$P(k) \propto k^{-3} \quad (7)$$

Assim se mostra que a distribuição de grau deste modelo de redes complexas (em que a única regra é a da preferência linear na escolha do nodo destinatário de cada nova ligação) reflecte uma estrutura *scale-free*, com expoente $\gamma = 3$ [4,5]. Este valor do expoente é exacto [6]. A expressão (7) é um resultado válido para redes de tamanho muito elevado e apenas em determinados intervalos ($k \gg 1$), existe um $k_{cut\ off}$ que dá uma medida do grau máximo presente na rede, e que é determinado pelos efeitos de tamanho finito.

Convém enfatizar que o princípio da ligação preferencial é a ideia base de toda a teoria de redes evolutivas moderna. Através deste princípio podem ser geradas redes *scale-free* com expoentes γ numa gama $]2, \infty[$, apenas à custa de uma pequena alteração nas regras de preferência, como se mostrará na próxima secção. Expoentes inferiores a 2 reflectem uma estrutura em que o grau médio diverge, o que só é possível em redes cujo número de arestas cresce mais depressa que o número de vértices. Por sua vez um expoente $\gamma = \infty$ corresponde a uma distribuição de grau exponencial que é equivalente ao caso em que não há preferência, os vértices antigos são escolhidos aleatoriamente com igual probabilidade para todos [1].

1.1.2-APROXIMAÇÃO AO CONTÍNUO

A aproximação ao contínuo foi usada na secção anterior para obter o valor exacto do expoente que caracteriza a distribuição de grau do modelo B-A, $\gamma = 3$. Os primeiros expoentes γ para o modelo B-A, descrito anteriormente, foram obtidos através desta abordagem [4]. O emprego desta técnica conduz aos valores exactos dos expoentes em numerosos modelos de redes evolutivas *scale-free* e permite descrever as principais características estruturais com facilidade [1,7,8]. Para descrever a técnica olhar-se-á novamente para o modelo B-A em concreto, por fornecer um exemplo simples e esclarecedor.

No modelo B-A a única dinâmica é a entrada, em cada iteração, de um novo nó na rede que estabelece uma ligação com outro previamente existente, seguindo a regra da preferência linear. Relembrando a equação (2):

$$p(k, i, t+1) = \frac{k-1}{2t} p(k-1, i, t) + \left(1 - \frac{k}{2t}\right) p(k, i, t)$$

com condição inicial $p(k, i=0, 1, t=1) = \delta_{k,1}$ e condição fronteira $p(k, i=t, t) = \delta_{k,1}$. A equação (2) pode ser reescrita da seguinte forma:

$$2t[p(k, i, t+1) - p(k, i, t)] = (k-1)p(k-1, i, t) - kp(k, i, t)$$

passando t e k ao limite contínuo tem-se:

$$2t \frac{\partial p(k, i, t)}{\partial t} + \frac{\partial [kp(k, i, t)]}{\partial k} = 0 \quad (8)$$

Aplicando $\int_0^\infty dk$ à equação anterior, e integrando por partes, obtém-se:

$$\frac{\partial k(i, t)}{\partial t} = \frac{k(i, t)}{2t} \Leftrightarrow \frac{\partial k(i, t)}{\partial t} = \frac{k(i, t)}{\int_0^t duk(i, t)} \quad (9)$$

onde se faz uso da igualdade $\int_0^t dk(i,t) = 2t$. $k(i,t)$ representa o grau médio do nó i no tempo t sobre as possíveis configurações da rede, $k(i,t) = \int_0^\infty dk k p(k,i,t)$. A expressão (9) é chamada equação de taxa do grau médio. A sua condição inicial é $k(0,1;t=1)=1$ e a condição fronteira é $k(t,t)=1$. A solução geral de equação (9) é da forma:

$$k(i,t) = C(i)t^{1/2} \quad (10)$$

onde $C(i)$ é uma função arbitrária de i . Utilizando a condição fronteira $k(t,t)=1$, vê-se que a equação (9) toma a forma:

$$k(i,t) = \left(\frac{t}{i}\right)^{1/2} \quad (11)$$

Encontra-se um comportamento *scale-free* na função do grau médio, com expoente $\beta = 1/2$, como é sobejamente sabido existir no modelo B-A.

Na aproximação ao contínuo a distribuição de grau é dada por:

$$P(k,t) = \frac{1}{t} \int_0^t di \delta(k - k(i,t)) = -\frac{1}{t} \left(\frac{\partial k(i,t)}{\partial i} \right)^{-1}_{[i:k(i,t)=k]} \quad (12)$$

Introduzindo nesta expressão a equação (11) rapidamente se percebe que a relação entre os expoentes γ e β é:

$$\gamma = 1 + \frac{1}{\beta} \quad (13)$$

No caso do modelo B-A o expoente γ fica igual a 3, como se viu anteriormente, e é referido em inúmeros documentos científicos.

Introduzindo uma pequena modificação na regra de preferência pode-se fazer variar o expoente γ no intervalo $]2, \infty[$. Agora considera-se uma preferência não linear, em que a probabilidade de um vértice antigo ser o destinatário de cada uma das novas ligações é proporcional ao seu grau somado a uma constante. Neste modelo cada novo vértice estabelece m ligações com os vértices preexistentes. Ou seja, na equação (9) substituímos o numerador do lado direito pela multiplicação de m pela soma do grau $\bar{k}(i, t)$ com uma constante A , chamada atracção adicional. Aqui quando um novo vértice chega à rede liga-se a m vértices antigos seguindo o princípio da ligação preferencial. Reescrevendo (9) tem-se:

$$\frac{\partial k(i, t)}{\partial t} = m \frac{k(i, t) + A}{\int_0^t du (k(i, u) + A)} = m \frac{k(i, t) + A}{t(2 + A)} \quad (14)$$

Cuja solução geral é da forma:

$$k(i, t) = C(i) t^{\frac{1}{2+A}}$$

A condição fronteira $k(t, t) = m$ implica que $C(i) = m \cdot s^{-\frac{1}{2+A}}$, e a solução vem:

$$k(i, t) = m \left(\frac{t}{s} \right)^{\frac{1}{2+A}} \quad (15)$$

Como podemos ver o comportamento *scale-free* mantém-se, mas aqui o expoente $\beta = (2 + A)^{-1}$ depende de A . Facilmente se percebe que a atracção adicional pode tomar qualquer valor entre no intervalo $] -1, \infty[$, um A menor que -1 corresponde a um $\beta > 1$, o que significaria que o grau do vértice i cresce mais rapidamente que o número de vértices na rede, o que não é possível nas redes aqui consideradas, pois o número de arestas criado por unidade de tempo não cresce e é constante.

No limite de k e t contínuos tem-se:

$$P(k, t) = \frac{1}{t} \int_0^t di \delta(k - k(i, t)) = -\frac{1}{t} \left(\frac{\partial k(i, t)}{\partial i} \right)^{-1}_{[i: k(i, t) = k]} \quad (16)$$

Considerando uma distribuição de grau estacionária obtemos a relação bem conhecida:

$$\gamma = 1 + \frac{1}{\beta} = 3 + A \quad (17)$$

Quando $A = 0$ a preferência é linear com o grau, correspondendo ao modelo B-A com $\gamma = 3$; quando $A \rightarrow -1$ o expoente $\gamma \rightarrow 2$ que é o mínimo valor possível para estas redes. Assim variando A tem-se acesso a toda a gama de $\gamma \in]2, \infty[$.

A técnica da aproximação ao contínuo, não obstante de ser muito simples e útil tem algumas limitações relacionadas com a exactidão da aproximação. No limite $t \rightarrow \infty$ a função delta de Dirac descreve com exactidão a probabilidade de um vértice nascido num determinado tempo ter um certo grau. Na realidade, para um t finito, a probabilidade de um vértice ter um determinado grau é uma função de Poisson, que é semelhante a um delta de Dirac mas tem largura não nula. Por isto, a aplicabilidade desta técnica limita-se a sistemas muito grandes.

1.1.3-CAMINHOS E DISTÂNCIA MAIS CURTA ENTRE VÉRTICES

Num grafo G , ao conjunto de todos os vértices chama-se V , e ao conjunto de todas as arestas E . Define-se como caminho entre quaisquer dois vértices $i, j \in V$ qualquer conjunto de vértices e arestas que permitam chegar do vértice i ao j , viajando de i para um dos seus vizinhos, v , através de aresta $\{i, v\}$, do vizinho v para um vizinho de v , e assim consecutivamente até atingir j . De todos esses caminhos o mais curto é aquele ou aqueles que minimizar a distância, l_{ij} . Em grafos não pesados todas as arestas têm o mesmo peso, portanto considera-se que o comprimento é igual para todas as arestas e é unitário. O caminho mais curto é aquele que envolver menor número de arestas. A distância do caminho mais curto l_{ij} é o número de arestas percorridas no caminho considerado, e é igual a l_{ji} quando as arestas não são direccionadas. Na generalidade é possível que existam vários caminhos mais curtos de igual comprimento.

A distância média dos caminhos mais curtos sobre todos os pares de vértices e todas as realizações da rede é também chamado o diâmetro, e determina o “tamanho linear” efectivo da rede.

Os caminhos e distâncias mais curtas são a base de muitas medidas da centralidade dos vértices na rede, isto é, da sua importância nas comunicações entre os vértices de rede. Várias medidas diferentes de centralidade aludem a noções diferentes de importância na rede. O tema deste estudo incide precisamente numa destas medidas, a Centralidade de Intermediação (*Betweenness Centrality*), neste texto é referida usando as iniciais do nome inglês BC.

1.1.4-ÁRVORES

As árvores são redes em que não existem circuitos fechados ou *loops*, isto é, existe um só caminho entre todos os pares de vértices. Da perspectiva das redes evolutivas é equivalente a uma dinâmica em que cada novo nodo se liga a um só dos vértices preexistentes, e o grau médio $\langle k \rangle = 2$. Como facilmente se pode constatar esta condição é necessária e suficiente para que numa rede evolutiva não existam *loops*. De resto, todas as deduções e expressões encontradas anteriormente são válidas para estes grafos. O próprio modelo B-A gera redes com esta característica.

Apesar de não existirem diferenças no tratamento destas redes em particular, no que respeita ao cálculo do grau médio $k(v, t)$ e à distribuição de grau $P(k)$, existem algumas diferenças estruturais fundamentais nestas redes. Por exemplo, enquanto em geral as redes *scale-free* têm uma distância média dos caminhos mais curtos que cresce com o logaritmo do tamanho (N), nas árvores cresce com $N^{1/2}$ se $\gamma > 3$ e com $N^{\frac{\gamma-2}{\gamma-1}}$ se $2 < \gamma < 3$ [1].

A presença de *loops*, que multiplicam o número de caminhos entre pares de nodos, faz com que existam vários caminhos mais curtos de igual comprimento em simultâneo. No caso de existir um só caminho entre todos os pares de vértices, cada nodo v integrante da rede “vê”, da sua posição, $k(v)$ zonas isoladas, cujas comunicações internas nunca o envolvem, mas que precisam dele para todas as comunicações com as restantes zonas, como se pode observar na Figura 1. Por outras palavras, se o vértice v fosse apagado da rede esta ficaria dividida em $k(v)$ zonas completamente isoladas. Esta característica faz com que o nodo v seja indispensável à comunicação entre as diversas $k(v)$ zonas. Isto é uma característica exclusiva das árvores, por isso pode antecipar-se alguma particularidade também no que respeita à centralidade. É principalmente sobre este tipo de grafos, árvores, que este trabalho incidirá.

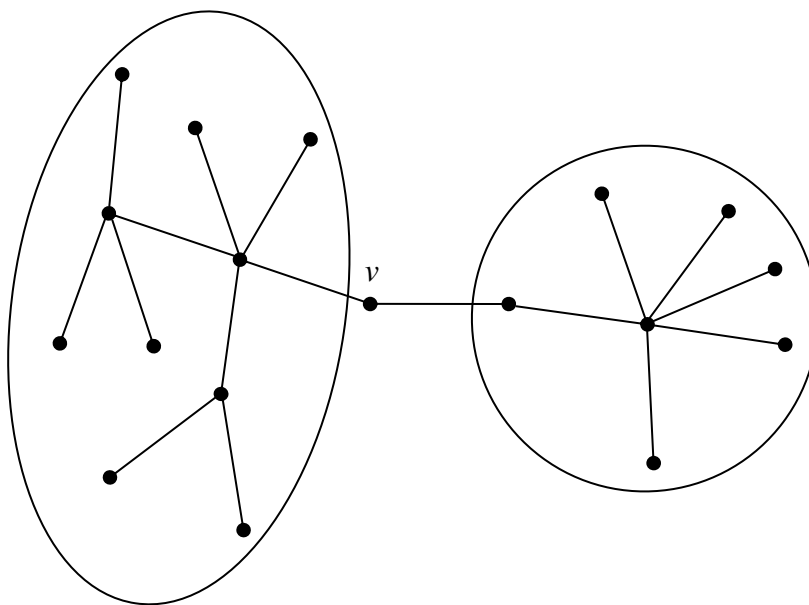


Figura 1: Representação esquemática da estrutura de um grafo tipo árvore. O vértice v liga duas regiões da rede que na sua ausência e das suas arestas estariam completamente desligadas. Esta propriedade é uma característica exclusiva das árvores.

1.2-MEDIÇÃO DA CENTRALIDADE

À partida, pode pensar-se que uma boa medida da centralidade de um nodo seria o seu grau, mas na realidade a questão da centralidade é mais complexa. Olhe-se para o exemplo da Figura 2, aqui o nodo v tem um grau baixo, apenas duas arestas, e desempenha um papel de elevada centralidade. Esta discrepância explica-se comparando o carácter local da conectividade com o carácter global da centralidade, que abrange a importância do nodo para toda a rede.

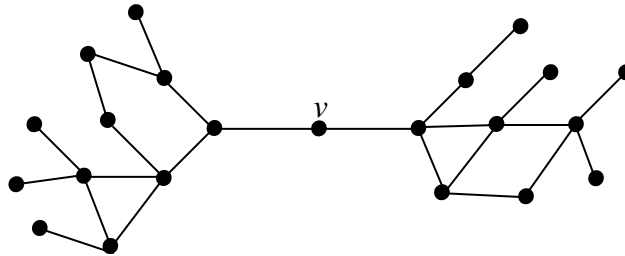


Figura 2: Neste esquema o nodo v tem apenas duas ligações, ou seja a sua conectividade é baixa. Por outro lado a sua centralidade é muito elevada, uma vez que une dois grandes componentes que na sua ausência estariam completamente desligados.

O conceito de centralidade existe sob a forma de algumas variações, todas tentam reflectir a noção de importância de um vértice num grafo. Denomina-se o número de caminhos mais curtos entre quaisquer nodos $s, t \in V$ como σ_{st} , naturalmente $\sigma_{ts} = \sigma_{st}$, e $\sigma_{ss} = 1$ por convenção. O número de caminhos mais curtos entre os nodos $s, t \in V$ que passam por v denota-se $\sigma_{ij}(v)$. As seguintes são algumas das medidas da centralidade mais habituais:

$$C_C(v) = \frac{1}{\sum_{s \in V} l_{sv}} \quad \text{Closeness Centrality (Sabidussi, 1966)}$$

$$C_G(v) = \frac{1}{\max_{s \in V} l_{sv}} \quad \text{Graph Centrality (Hage e Harary, 1995)}$$

$$C_S(v) = \sum_{s \neq v \in V} \sum_{t \neq s, v \in V} \sigma_{sv}(v) \quad \text{Stress Centrality (Shimbel, 1953)}$$

$$C_B(v) = \sum_{s \neq v \in V} \sum_{t \neq s, v \in V} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad \text{Betweenness Centrality (Freeman, 1977; Anthonisse, 1971)}$$

Geralmente elevados índices de centralidade indicam que o nodo está separado dos outros nodos da rede por caminhos mais curtos relativamente pequenos ou que está numa grande fracção dos caminhos mais curtos entre os outros nodos. Estes índices podem ser normalizados de forma a variarem no intervalo $[0,1]$.

1.2.1-CENTRALIDADE DE INTERMEDIACÃO (BC)

Uma das propriedades mais conhecidas das redes *scale-free* e o efeito *small-world*, que consiste num crescimento do diâmetro da rede mais lento que o crescimento do tamanho da própria rede. A distância média entre pares de vértices cresce com o logaritmo do número de vértices na rede [9] e um logaritmo cresce mais lentamente que qualquer lei de potência. O efeito *small-world* faz com que, mesmo para números de vértices muito grande, as distâncias dos caminhos mais curtos sejam pequenas. Nas redes *scale-free* acontece que um pequeno número nodos tem um grau elevado, agindo como *hubs*, fazem a ligação entre muitos nodos de grau menor, e tem um papel determinante na redução de diâmetro do sistema.

Muitas redes complexas reais possuem esta propriedade, o pequeno diâmetro permite que os diversos agentes comuniquem ou interajam com outros rapidamente usando os caminhos mais curtos. Os caminhos mais curtos são efectivamente de grande relevância no que toca às propriedades de transporte. Quando um pacote de dados é enviado de um vértice para outro numa rede *scale-free*, tal como a Internet, é mais eficiente que adopte o caminho mais curto entre os vértices. Os vértices com grande grau acabam por ficar muito “carregados” e “entupidos” devido ao grande número de pacotes que passam por ele ao longo dos caminhos mais curtos. Para evitar este tipo de congestões de tráfego e permitir aos pacotes viajarem num estado de “fluxo livre”, tem de se ampliar a capacidade de transmissão de dados de cada vértice até ser suficientemente grande para suportar a carga de transmissões definida pela sua posição na estrutura.

Esta medida da centralidade, tal como foi definida na secção anterior, conta a fracção de caminhos mais curtos entre todos os pares que passa por v . Na verdade o conceito de Centralidade de Intermediação (BC) surgiu no contexto das redes sociais para quantificar o poder centralizado nas pessoas numa rede social [10,11]. A partir daqui o símbolo usado para denotar a BC de v será $g(v)$. A definição de $g(v)$, fornecida na secção anterior, vem:

$$g(v) = \sum_{s \neq v \in V} \sum_{t \neq s, v \in V} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (18)$$

1.2.2-FUNÇÃO DISTRIBUIÇÃO DE BC

Para caso de redes *scale-free* foi apresentado um estudo numérico por Goh *et al* da distribuição de BC para modelos de redes *scale-free* estáticos em [12]. Nesse modelo $\gamma \in]2, \infty[$ é um parâmetro ajustável. Também estudaram a distribuição de BC no modelo B-A, previamente descrito neste texto, em que $\gamma = 3$. Eles mostram que a distribuição de BC segue uma lei de potência:

$$P(g) \sim g^{-\delta} \quad (19)$$

Este comportamento observa-se para g elevados até um g_{cutoff} , análogo a k_{cutoff} , determinado pelos efeitos de tamanho finito. Nesse estudo concluíram que o valor de $\delta \approx 2,2$ é universal para os valores de $\gamma \in]2, 3[$. Posteriormente, numa publicação dos mesmos autores em que são analisadas diversas redes reais e *in-silico* [13], argumenta-se que a distribuição de BC tem a forma da expressão (19) em todos os sistemas *scale-free* estáticos e evolutivos. Além disso, concluem que existem duas classes de redes *scale-free*, nas quais o valor de δ é universal. Concretamente na Classe I $\delta \approx 2,2$, enquanto na Classe II $\delta \approx 2,0$. δ pode ser calculado analiticamente para árvores, fornecendo um valor exacto de $\delta = 2$. Isto sugere que $\delta = 2$ é o resultado exacto para a classe II.

Gerou-se uma discussão, entre os autores de [12,13] e Barthélemy autor de [14], centrada na afirmação de universalidade de δ . Barthélemy defende que o valor de δ não goza de universalidade. O presente trabalho pretende contribuir para o esclarecimento deste problema.

1.2.3-DEPENDÊNCIA COM A CONECTIVIDADE

A relação entre a BC e conectividade de um nodo foi estudada em [12,15], concluindo-se numericamente que é da forma:

$$g \sim k^\eta \quad (20)$$

onde o expoente η depende dos detalhes da rede. Esta relação entre as duas variáveis implica que a uma certa conectividade corresponde uma determinada BC. Devido aos efeitos de tamanho finito $g(k)$ está sujeita a pequenas flutuações. A já referida g_{cutoff} , a partir da qual a lei de potência não se ajusta aos dados, pode ser estimada a partir de (20) substituindo por k_{cutoff} .

Uma vez que g e k estão relacionados através de (20), usando a aproximação ao contínuo previamente discutida, podemos escrever

$$P(g) = \int dk P(k) \delta(g - k^\eta) \quad (21)$$

Podemos ver em (20) que um grande g implica um grande k , então podemos escrever:

$$P(g \gg 1) = \int dk k^{-\gamma} \delta(g - k^\eta) \quad (22)$$

Resolvendo a expressão anterior chega-se ao seguinte resultado:

$$P(g \gg 1) \sim g^{-1 - \frac{\gamma-1}{\eta}}$$

Que corrobora a igualdade encontrada, recorrendo a outras abordagens, em [12,15]:

$$\eta = \frac{\gamma-1}{\delta-1} \quad (23)$$

Agora existem condições para reescrever a expressão (20) à custa dos expoentes γ e δ :

$$g \sim k^{\frac{\gamma-1}{\delta-1}} \quad (24)$$

Esta expressão define o comportamento do expoente η . O único parâmetro do qual não se conhecem totalmente todos os aspectos é δ . Mas pode notar-se que se δ exibir universalidade η será uma função linear de γ com a dependência $(\delta-1)^{-1}$ (equação (23)).

1.3-MODELO COMPUTACIONAL

Foi analisado na secção 1.1.2 uma variação do modelo B-A, que e pode produzir estruturas *scale-free* com expoente γ num intervalo $]2, \infty[$. Nesse modelo a regra é que os novos vértices se ligam aos antigos com probabilidade proporcional à soma ao seu grau somado com a atracção adicional $A \in]-1, \infty[$.

O algoritmo utilizado para gerar as árvores evolutivas *scale-free*, com N vértices, que serão sujeitas a análise compõe-se dos seguintes passos:

- i) criar dois vértices interligados, com índices 0 e 1, e iniciar a variável $t = 2$,
- ii) criar um vértice com o índice $i = t$,
- iii) ligar o novo vértice i a um dos anteriores, j , escolhido aleatoriamente com probabilidade igual $k(j) + A$,
- iv) incrementar t e repetir os passos ii) e iii), enquanto $t < N$.

Este algoritmo, quando executado para um N suficientemente elevado, gera simulações de árvores *scale-free* com expoente $\gamma \approx 3 + A$ [6,17].

1.4-ALGORITMO PARA O CÁLCULO DA BC

O problema de calcular os índices de BC é um pouco mais complexo. O principal obstáculo é a limitação dos recursos de computação, principalmente a velocidade do processamento. Os algoritmos convencionais exigem um tempo de computação $\Theta(N^3)$ e ocupam uma memória $\Theta(N^2)$. Por cada nodo temos de calcular uma contribuição de cada par de nodos, o número de pares é aproximadamente N^2 para $N \gg 1$, uma vez que o número de nodos é N o número de operações cresce com N^3 . Estas condições são proibitivas ao cálculo da BC em redes com tamanhos elevados. Recentemente foi proposto por Brandes em [16] um algoritmo cumulativo para o cálculo dos índices de centralidade que reduz o tempo de computação a $\Omega(NM)$ e a memória necessária a $\Omega(N + M)$, sendo M o número de arestas.

Um vértice $v \in V$ encontra-se num caminho mais curto entre dois vértices $s, t \in V$ se e só se $l_{sv} + l_{vt} = l_{st}$. Então dado o número de caminhos mais curtos entre cada par e os seus respectivos comprimentos pode escrever-se o número de caminhos mais curtos entre $s, t \in V$ que passa por $v \in V$ desta forma:

$$\sigma_{st}(v) = \begin{cases} 0 & \text{se } l_{st} < l_{sv} + l_{vt} \\ \sigma_{sv} \cdot \sigma_{vt} & \text{se } l_{st} = l_{sv} + l_{vt} \end{cases} \quad (25)$$

A dependência da BC de $v \in V$ com o par $s, t \in V$ é $\lambda_{st}(v) = \frac{\sigma_{st}(v)}{\sigma_{st}}$, para obter a BC de

um nodo apenas temos de somar todos os $\lambda_{vt}(v)$:

$$g(v) = \sum_{s \neq v} \sum_{t \neq s, v} \lambda_{st}(v) \quad (26)$$

Resulta que os algoritmos convencionais fazem o cálculo da centralidade BC em dois passos. Primeiro calculam o número e o comprimento dos caminhos mais curtos entre todos os pares, de seguida somam os $\lambda_{vt}(v)$ calculados com a relação (25).

A matriz de adjacência de um grafo não pesado é uma matriz $B = (b_{ij})_{i, j \in V}$ com dimensões $N \times N$, onde $b_{ij} = 1$ se a aresta $\{i, j\} \in E$ ou $b_{ij} = 0$ caso contrário. Algumas

das implementações mais usadas na computação do índice de BC fazem uso da seguinte propriedade algébrica: Se $B^n = (b_{ij}^{(n)})_{i,j \in V}$ for a n -ésima potência da matriz de adjacência de um grafo não pesado então $b_{ij}^{(n)}$ é igual ao número de caminhos entre os nodos i e j com comprimento exactamente n .

Uma vez que para cada vértice é necessário somar N^2 termos o tempo necessário ao cálculo da BC é predominantemente gasto na multiplicação de matrizes. É óbvio que a contagem algébrica do número de caminhos entre pares computa muita informação supérflua. Só existe interesse nos caminhos mais curtos, e não em todos os caminhos de comprimento inferior ao mais longo de todos os caminhos mais curtos. O algoritmo BFS (*breadth-first search*) começa com um determinado vértice $s \in V$ e em cada iteração adiciona mais um vértice à sua lista de vértices descobertos, por ordem de proximidade. Para isso avança ao longo das arestas, naturalmente descobre todos os caminhos mais curtos do vértice s . Um predecessor do vértice $w \in V$, nos caminhos mais curtos de $s \in V$, é um vértice v vizinho de w tal que $l_{sw} = l_{sv} + 1$. Ou seja o conjunto dos predecessores de w nos caminhos mais curtos de s é:

$$P_s(w) = \{v \in V : \{v, w\} \in E, l_{sw} = l_{sv} + 1\} \quad (27)$$

Levando em conta que a última aresta de qualquer caminho mais curto entre s e w é sempre uma aresta $\{v, w\} \in E$, ou seja, é sempre entre w e um vizinho seu. A seguinte relação torna-se evidente.

$$\sigma_{sw} = \sum_{i \in P_s(w)} \sigma_{si} \quad (28)$$

Por uma questão de simplicidade elimina-se a soma explícita sobre todos os pares de vértices introduzindo a dependência da BC de um vértice $v \in V$ com um único vértice $s \in V$, definida desta forma:

$$\lambda_{s*}(v) = \sum_{t \in V} \lambda_{st}(v) \quad (29)$$

A inovação do algoritmo proposto por Brandes consiste em explorar a propriedade recursiva do somatório (29). É particularmente fácil reconhecer a recursividade no caso especial em que existe exactamente um caminho mais curto entre cada par de vértices.

Esta condição implica que os vértices e as arestas de todos os caminhos mais curtos de $s \in V$ formem uma árvore. Consequentemente, v está em todos ou em nenhum dos caminhos entre s e qualquer $t \in V$, isto é, $\lambda_{st}(v)$ é igual a um ou a zero. E ainda, v está em todos os caminhos mais curtos para os vértices de que é predecessor partindo de s , e em todos os caminhos mais curtos em que, por sua vez, estes estão (ver Figura 3). Expressando de outro modo:

$$\lambda_{s\bullet}(v) = \sum_{w:v \in P(w)} (1 + \lambda_{s\bullet}(w)) \quad (30)$$

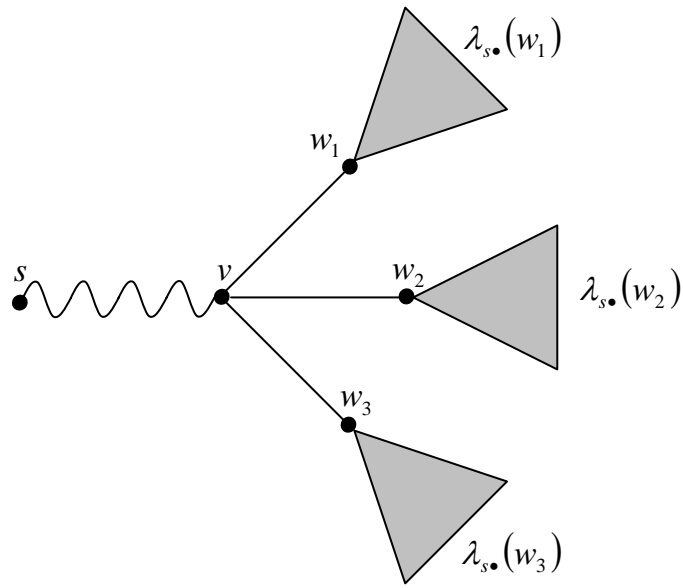


Figura 3: Vê-se nesta representação que v está em todos os caminhos mais curtos de s para os vértices de que v é predecessor (w_1 , w_2 e w_3), bem como todos os caminhos mais curtos em que, por sua vez, estes estão.

Os grafos descritos anteriormente obedecem a esta regra, de facto nas árvores existe um e um só caminho entre cada par de vértices.

No caso geral aparece uma relação semelhante, apesar de este estudo incidir sobre árvores fica aqui a expressão geral:

$$\lambda_{s\bullet}(v) = \sum_{w:v \in P(w)} \frac{\sigma_{sv}}{\sigma_{sw}} (1 + \lambda_{s\bullet}(w)) \quad (31)$$

Com base na expressão (30) foi implementado um algoritmo computacional, do qual se pode ver uma representação esquemática de seguida:

```

 $g[v] \leftarrow 0, v \in V$  ;
for  $s \in V$  do
     $S \leftarrow$  stack vazia;
     $P[w] \leftarrow$  espaço vazio,  $w \in V$  ;
     $l[t] \leftarrow -1, t \in V$  ;  $l[s] \leftarrow 0$  ;
     $Q \leftarrow$  fila vazia;
    colocar  $s$  na fila  $Q$  ;
    while  $Q$  não estiver vazia do;
        retirar o primeiro na fila  $v \leftarrow Q$  ;
        coloca-lo na stack  $v \rightarrow S$ 
        foreach vizinho  $w$  de  $v$  do
            // é a primeira vez que  $w$  é encontrado?
            if  $l[w] < 0$  then
                colocar  $w$  no último lugar da fila  $w \rightarrow Q$  ;
                 $l[w] \leftarrow l[v] + 1$  ;
                 $v \rightarrow P[w]$  ;
            end
        end
    end
    //  $S$  devolve os vértices por ordem não decrescente de distância a  $s$ 
    while  $S$  não estiver vazia do
        retirar o último da stack  $w \rightarrow S$ 
        if  $w \neq s$  then
             $v \leftarrow P[w]$ 
             $\lambda(v) \leftarrow \lambda(v) + \lambda(w) + 1$ 
             $g(w) \leftarrow g(w) + \lambda(w)$ 
        end
    end
end

```

Em [16] é apresentada uma versão muito semelhante a este algoritmo que faz uso da expressão (31) e é capaz de calcular os índice de BC em redes com múltiplos caminhos mais curtos entre pares de vértices, que pode ser facilmente adaptado a grafos pesados.

2-BC EM ÁRVORES *SCALE-FREE*

Foi mencionado previamente que nas árvores existe um único caminho entre cada par de nodos, e que o facto de não existirem caminhos alternativos nestas redes poderá fornecer-lhes alguma particularidade no que respeita aos índices de centralidade. Explicitamente, a BC é uma função exclusiva dos caminhos mais curtos, e o facto de estes serem únicos entre pares simplifica o cálculo deste índice. Também a estrutura *scale-free* tem influência na BC, pois estes grafos têm a propriedade *small-world*, a distância dos caminhos é muito pequena e cresce muito devagar com o número de vértices.

Durante as próximas secções tira-se partido da propriedades da estrutura *scale-free* e do facto de os caminhos serem únicos tendo em vista a descrição do comportamento da BC em árvores. Designadamente estuda-se a relação da BC com a conectividade dos vértices e a sua distribuição nestes grafos.

2.1-TRABALHO ANTERIOR

Em [13] Goh *et al* propuseram, baseando-se num estudo extensivo tanto de redes reais como de simulações computacionais, que existem duas classes distintas de redes complexas no que respeita à BC. Segundo esse trabalho, dentro de cada classe o expoente da função distribuição de BC é universal, para a classe I $\delta \approx 2,2$, e para a classe II $\delta \approx 2,0$. O caso das árvores *scale-free* integra-se, segundo esta distinção, na classe II.

O expoente da distribuição de BC é calculado analiticamente, para a variação do modelo B-A proposto na secção 1.1.2, por Goh *et al* em [13].

Na verdade a BC considerada por Goh *et al* não é a dos vértices, mas sim a das arestas. O conceito é o mesmo, com a substituição do número de caminhos mais curtos que passa num vértice pelo número de caminhos mais curtos que passa por uma aresta. Os autores defendem que nas árvores a BC dos nodos e das arestas manifestam o mesmo comportamento, baseando esta afirmação apenas nos resultados das simulações. A partir de agora a BC das arestas escrever-se-á g_e , e a dos vértices g_v .

Considere-se a variação do modelo B-A já referida e adopte-se $m = 1$. Cada aresta “vê” dois componentes da rede ligadas apenas por si, e a soma do número de vértices das duas regiões é o número total de vértices na rede. Chame-se $P(d;e,t)$ à probabilidade de a aresta nascida no tempo e ligar um componente com $t+1-d$ vértices do lado dos vértices ascendentes (ou seja, do lado dos vértices mais antigos que e) e d vértices do lado dos descendentes (ou seja, do lado onde todos são mais jovens que e) no tempo t . Da estrutura das árvores resulta que a BC das arestas é igual a duas vezes o produto entre os dois componentes, $g_e = 2d(t+1-d)$. A equação de mestra para a probabilidade $P_e(d;e,t)$ é esta:

$$P(d;e,t+1) = r_1(d,t)P(d;e,t) + r_2(d-1,t)P(d-1;e,t) \quad (32)$$

onde $r_1(d,t)$ representa a probabilidade do o vértice criado em $t+1$ se ligar ao componente com $t+1-d$ vértices do lado dos ascendentes, e $r_2(d-1,t)$ a probabilidade de se ligar ao lado com $d-1$ vértices do lado dos descendentes. Na sua forma explicita:

$$r_2(d,t) = 1 - r_1(d,t) = \frac{d(2+A)-1}{2t+A(t+1)} \quad (33)$$

recorde-se que entre $t+1$ nodos existem t arestas. Se introduzirmos a média sobre todas as arestas, que é a probabilidade de uma aresta ao acaso ter dois componentes com d e $t+1-d$ vértices:

$$P(d, t) = \frac{1}{t} \sum_{e=1}^t P(d; e, t) \quad (34)$$

Aplicando $\sum_{e=1}^t$ aos dois lados da equação (32) e escrevendo em termos de $P(d, t)$ aparece:

$$(t+1)P(d, t+1) = r_1(d, t)tP(d, t) + r_2(d-1, t)tP(d-1, t) \quad (35)$$

No limite $t \rightarrow \infty$, e escrevendo $P(d, t)$ na forma de *scaling* $P(d/t) = P(x)$, a equação anterior vêm:

$$(t+1)P(x) - tP(x) \approx -x \frac{\partial P(x)}{\partial x} - P(x) \quad (36)$$

onde $x = m/t$, e foi usada a aproximação $P(x - 1/t) = P(x) - \frac{\partial P(x)}{t \partial x}$.

Daqui sai a relação:

$$\frac{\partial P(x)}{\partial x} = \frac{-2P(x)}{x}$$

cuja solução é:

$$P(x) \sim \frac{1}{x^2} \quad (37)$$

Podemos ver como $P(x)$ é independente da atracção adicional A . Usando a forma da BC para este caso $g_e = 2d(t+1-d)$ para t muito elevado e d finito tem-se $g_e = 2t^2 x$. Agora usando a equação (37) e a relação linear entre g e x podemos escrever:

$$P(g) \sim \frac{1}{g_e^2} \quad (38)$$

Em redes evolutivas que crescem segundo estas regras o expoente δ_e da distribuição de BC das arestas, é exactamente igual a 2, e é independente de γ ou qualquer outro parâmetro.

2.2-EQUAÇÕES DIFERENCIAIS

Neste trabalho o problema da centralidade em árvores evolutivas *scale-free* é tratado com uma abordagem similar à da secção 1.1.2. Usa-se a aproximação ao contínuo e escrevem-se equações diferenciais. Como foi referido anteriormente, numa árvore só existe um caminho entre cada par de nodos, consequentemente cada nodo v une um número, igual ao seu grau $k(v)$, de regiões da rede ligadas apenas por si próprio, Figura 4.

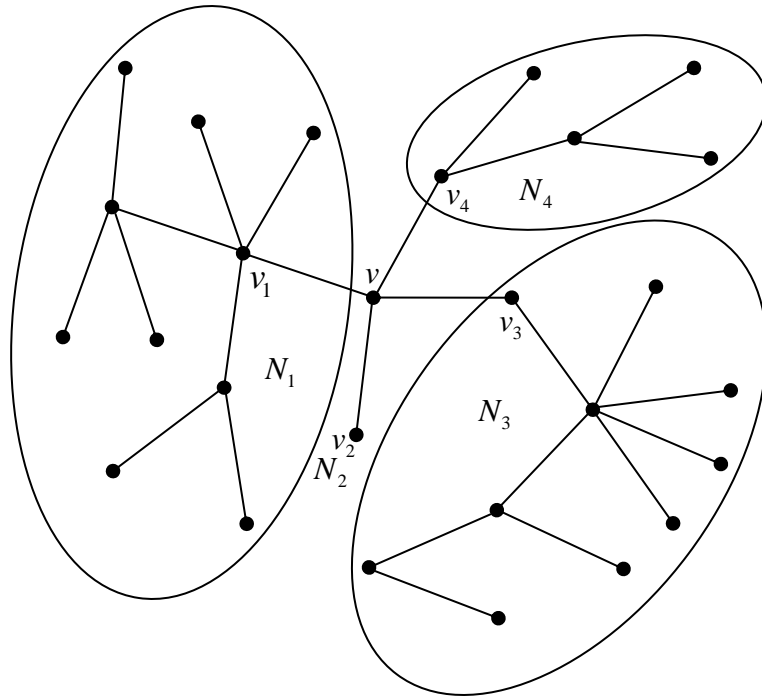


Figura 4: Representação esquemática de um grafo sem *loops*. O nodo v , da sua perspectiva, vê um número igual ao seu grau de regiões apenas ligadas por si. Tal significa que v se encontra em todos os caminhos entre nodos de regiões diferentes.

Nestas condições o índice BC no nodo v é igual ao somatório da multiplicação do número de vértices N_i de cada região $i \in \{1, \dots, k(v)\}$ pelo número de vértices de cada uma das restantes regiões. Isto é, no caso das árvores a definição de BC vem simplificada:

$$g_n(v, t) = \sum_{i=1}^{k(v, t)} \sum_{j(\neq i)=1}^{k(v, t)} N_i(v, t) N_j(v, t) \quad (39)$$

Pode escrever-se as equações para número de nodos $N_i(v, t)$ em cada uma destas regiões do vértice v no tempo t . Uma vez que, em cada passo de tempo, a probabilidade do novo vértice se ligar a uma destas regiões é igual à soma das probabilidades dos seus vértices receberem a ligação, pode-se escrever a variação de N_i no tempo deste modo:

$$\frac{\partial N_i(v, t)}{\partial t} = \frac{N_i(v, t)(2 + A) - 1}{2t + A(t + 1)} \quad (40)$$

Para as $k(v, t)$ equações (40) existem duas condições fronteiras. Para a região N_1 , dos ascendentes do vértice, a condição fronteira é diferente das outras:

$$N_1(v, t = v) = v \quad (41)$$

$$N_{i>1}(v, t : k(v, t) = i) = 1$$

Lembrando que nas árvores *scale-free* o grau do nodo v no tempo t vem $k(i, t) = \left(\frac{t}{i}\right)^\beta$, chega-se às soluções dos N_i 's:

$$N_1(v, t) = \left(v - \frac{1}{2 + A}\right) \frac{t}{v} + \frac{1}{2 + A}$$

$$N_{i>1}(v, t) = \left(1 - \frac{1}{2 + A}\right) i^{-1/\beta} \frac{t}{v} + \frac{1}{2 + A}$$

para $t \rightarrow \infty$. Nas condições $1 \ll v \ll t$, as expressões anteriores reduzem-se a:

$$N_1(v, t) \approx t \quad (42)$$

$$N_{i>1}(v, t) \approx \left(1 - \frac{1}{2 + A}\right) i^{-1/\beta} \frac{t}{v}$$

As soluções (42) mostram que a região dos ascendentes cresce linearmente com o tempo, absorvendo quase todos os vértices e arestas que chegam para se juntar à rede. Funciona de uma forma semelhante a um nodo de condensação [1,8], onde o que desempenha o papel da elevada atracção é o número de nodos v com que começa, como se pode ver em (41). Por outro lado as soluções para $i > 1$ exibem comportamento *scaling*.

2.3-CÁLCULO DA BC

Como se pode facilmente constatar da observação das expressões (42), a região onde estão os vértices e arestas mais velhos que v é muito maior que qualquer uma das outras, $N_1(v, t) \gg N_{i>1}(v, t)$. Já se referiu que do ponto de vista da aresta nascida no tempo v o lado dos ascendentes é muito maior que o dos descendentes. Isto está patente na primeira das expressões (42).

Estas circunstâncias permitem simplificar a relação (39):

$$g_n(v, t) \approx 2N_1(v, t) \sum_{i=2}^{k(v, t)} N_i(v, t) \quad (43)$$

uma vez que os produtos $N_{i \neq 1}(v, t)N_{j \neq 1}(v, t)$ são muito menores que os $N_1(v, t)N_{i \neq 1}(v, t)$ apenas se consideram os últimos. Utilizando de novo a relação $k(i, t) = \left(\frac{t}{i}\right)^\beta$, aparece a seguinte expressão para a BC dos nodos:

$$g_n(v, t) \approx \left(2 - \frac{2}{2 + A}\right) t k^{\frac{1}{\beta}} \sum_{i=2}^k i^{-\frac{1}{\beta}} \quad (44)$$

onde k é o grau do nodo v . Considerando que $\beta < 1$, como foi dito na secção 1.1.2, temos que o somatório da relação (44) é convergente mesmo para $k \rightarrow \infty$. Daí alcança-se a próxima relação entre a BC e a conectividade dos nodos nestas estruturas:

$$g \sim k^{\frac{1}{\beta} - \gamma - 1} \quad (45)$$

Este resultado exprime uma dependência linear entre a BC e a conectividade apenas quando a estrutura tem um expoente *scale-free* $\gamma = 2$. Quanto maior for γ mais forte é a dependência da BC com o grau. O que indica que apesar de para γ maior termos uma distribuição de grau mais igualitária entre vértices a distribuição de BC continua a ser muito desigual. A razão para este comportamento é que mesmo para uma distribuição de grau exponencial ($\gamma \rightarrow \infty$), os vértices mais antigos têm lugares mais centrais na estrutura apenas por terem nascido primeiro.

2.4-DISTRIBUIÇÃO DE BC

O resultado (45) pode ser usado para calcular a distribuição de BC dos nodos $P(g_n)$. Para isso emprega-se a técnica da aproximação ao regime contínuo já introduzida na secção 1.1.2, desta vez com uma pequena modificação:

$$P(g_n) = \frac{1}{t} \int dk P(k) \delta(g_n - g_n(k)) \quad (46)$$

para $k, g \gg 1$ a expressão anterior combinada com as relações (1) e (45) conduz ao seguinte resultado:

$$P(g_n) \sim \frac{1}{t} \int dk k^{-\gamma} \delta(g_n - k^{\gamma-1}) \sim \frac{1}{t} \left(k^{-\gamma} \left(\frac{\partial k^{\gamma-1}}{\partial k} \right)^{-1} \right)_{(k: k^{\gamma-1} = g_n)} \quad (47)$$

Por fim, resolvendo a equação anterior, surge uma lei de potência com expoente $\delta = 2$ para a distribuição de BC:

$$P(g_n) \sim \frac{1}{g_n^2} \quad (48)$$

Este resultado que se refere à BC dos nodos é igual ao encontrado por Goh *et al* em [13]. Apesar de, esse resultado se referir à BC das arestas é sugerido numericamente que ambos se comportam da mesma maneira.

Como foi dito na secção anterior, das equações (42) conclui-se que a região dos ascendentes tem um tamanho muito superior à soma dos tamanhos de todas as outras. Daí se depreende que a centralidade da aresta que nasceu em simultâneo com o nodo determina a sua centralidade.

Podemos relacionar g_n e g_e da seguinte forma:

$$g_n(v) = \frac{1}{2} \sum_{i=1}^k g_e(e_i, v) \quad (49)$$

onde $g_e(e_i, v)$ representa a BC da aresta ligada a v nascida em i -ésimo lugar da perspectiva de v . Representando o número de vértices do lado dos ascendentes da aresta

e_i do vértice v por $N_a(e_i, v)$, e o número do lado dos descendentes por $N_d(e_i, v)$, e lembrando que $g_e(e_i, v) = 2N_a(e_i, v)N_d(e_i, v)$ a equação (49) pode ser reescrita:

$$g_n(v) = \sum_{i=1}^k N_a(e_i, v)N_d(e_i, v) \quad (50)$$

Uma vez que para $v \gg 1$ $N_a(e_1, v) \approx N$, e o número de ascendentes das outras arestas $e_{i>1}$ de v é a soma de $N_a(e_1, v)$ com os $N_d(e_j, v)$ para $j > 1$ e $j \neq i$, como se vê no exemplo da Figura 5, é imediato que:

$$g_n(v) \approx N \sum_{i=1}^k N_d(e_i, v) \quad (51)$$

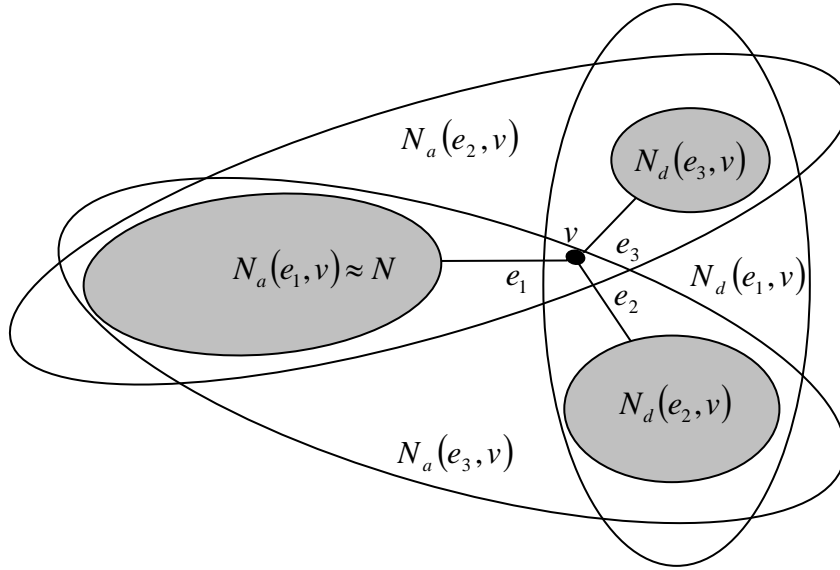


Figura 5: Pode observar-se neste esquema que o número de ascendentes de $e_{i>1}$ é igual ao número de ascendentes de e_1 somado com o número de descendentes de todas as outras arestas excepto e_i e e_1 . Se $N_a(e_1, v) \approx N$ então $N_a(e_i, v) \approx N$. Observa-se também a propriedade das árvores evolutivas de a soma de todos $N_d(e_{i>1}, v)$ ser igual a $N_d(e_1, v)$.

Na figura 5 também se vê a propriedade de o somatório em (51) igualar o dobro de $N_d(e_1, v)$, isto é $\sum_{i=2}^k N_d(e_i, v) = N_d(e_1, v)$. Donde:

$$g_n(v) \approx 2NN_d(e_1, v) \quad (52)$$

O termo da direita da expressão (52) é a BC da aresta criada em simultâneo com o nodo v , do que se deduz:

$$g_n(v) \approx g_e(e_1, v) \quad (53)$$

Como se tinha afirmado a BC de um nodo é dominada pela BC da aresta que nasceu com ele. Dado que por cada nodo existe exactamente uma aresta com a mesma idade as distribuições de BC dos nodos e das arestas são idênticas:

$$P(g_n) \approx P(g_v) \quad (54)$$

Este resultado justifica e confirma a assunção feita em [13], com efeito, mostra-se analiticamente que as duas distribuições de BC, dos nodos e das arestas, são aproximadamente iguais.

3-SIMULAÇÕES E ESTATÍSTICA

Com o fim de comprovar numericamente as relações obtidas nos capítulos anteriores procedeu-se à execução de simulações computacionais das redes aqui estudadas, para o que se utilizou o algoritmo introduzido na secção 1.3. Uma vez atingido o um determinado número de nodos N a informação gerada é analisada pelo algoritmo introduzido em 1.4 calculando a BC de cada vértice. As quantidades devolvidas neste processo são a BC e o grau de cada nodo, que depois são tratadas estatisticamente, e cujos resultados são aqui apresentados.

A distribuição de grau de uma rede *scale-free*, por definição, é de forma $P(k) \sim k^{-\gamma}$, $P(k)$ é a fracção de vértices que têm grau k . Pode-se calcular a fracção de vértices com grau igual ou superior a k usando a aproximação ao contínuo:

$$F(k) = \int_k^{\infty} dk' P(k') \sim k^{1-\gamma} \quad (55)$$

Logo a distribuição cumulativa $F(k)$ comporta-se como uma lei de potência com expoente $1-\gamma$. Tem a vantagem de apresentar muito menos ruído na análise de dados, como pode ser observado na Figura 6:

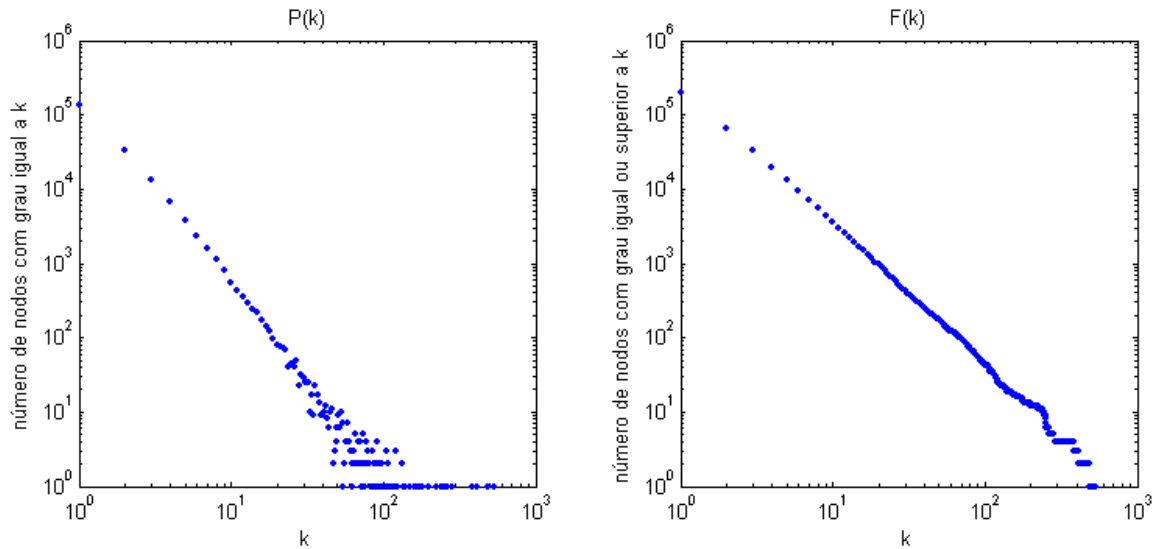


Figura 6: À direita a distribuição de grau $P(k)$ e à esquerda a distribuição cumulativa $F(k)$ numa árvore com $N = 200\,000$ e atracção adicional $A = 0$, ou seja, o modelo BA. Pode ver-se que o ruído é muito menor na distribuição cumulativa $F(k)$.

A representação logarítmica duma lei de potência é uma recta de declive igual ao expoente de lei de potência. Pelo que se pode estimar o expoente γ fazendo o ajuste duma recta aos pontos da representação logarítmica, segundo o método dos mínimos desvios quadrados. O expoente γ é calculado directamente do declive da recta de ajuste de acordo expressão (55).

No caso da distribuição de centralidade, $P(g) \sim g^{-\delta}$, o mesmo princípio pode ser aplicado, com o mesmo efeito, redução do ruído. Uma vez que a BC pode assumir valores até $\approx N^2$ a apenas existem N nodos na rede é necessário fazer um histograma do número de nodos que têm índice de centralidade g dentro de um certo intervalo em função do valor de g no centro do intervalo. Para observar o expoente δ constrói-se depois um histograma cumulativo de acordo com a técnica explicada acima para a distribuição de grau. A distribuição de centralidade cumulativa será semelhante à expressão (55):

$$F(g) = \int_g^{\infty} dg' P(g') \sim g^{1-\delta} \quad (56)$$

Para a estimativa do expoente δ faz-se o ajuste duma recta aos pontos da representação logarítmica, e calcula-se directamente através do declive da recta de ajuste de acordo expressão (56).

Quanto à relação entre a BC e a conectividade, $g \sim k^{\eta}$ usa-se uma técnica semelhante, a diferença está nos limites de integração:

$$G(k) = \int_0^{k^{\eta}} dg' g' \sim \int_0^k k'^{\eta} \sim k^{\eta+1} \quad (57)$$

A estimativa do expoente η faz-se da mesma forma que as dos outros expoentes, usando a representação logarítmica para calcular o declive da recta de ajuste, e daí η segundo (57).

Este método prova-se muito eficaz na redução de ruído, e é por isso muito popular no tratamento deste tipo de dados. Levaram-se a cabo simulações, de acordo com o modelo

adoptado para este estudo, com diversas atracções adicionais A , originando árvores *scale-free* com diversos expoentes $\gamma \approx 3 + A$.

A validade da aproximação (43) é robusta, como se pode observar na Figura 7, onde se mostra em simultâneo, para três estruturas com diferentes γ 's, a centralidade média e o produto $2 \cdot N_1(N - 1 - N_1)$ médio em função do grau. Estas médias são feitas em cada estrutura sobre todos os nodos com o mesmo grau. Como se pode ver a aproximação adequa-se quase perfeitamente.

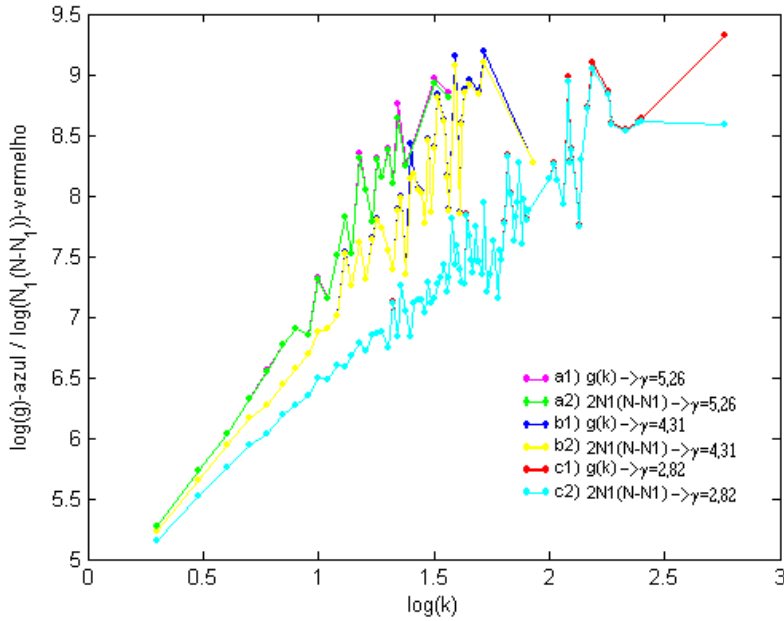


Figura 7: Três árvores (a, b e c) com diferentes $\gamma = \{5,26; 4,31; 2,82\}$. Para cada uma desenham-se as médias, sobre todos os vértices com o mesmo grau, da *betweenness* (1) e do produto $2 \cdot N_1(N - N_1)$ (2). Podemos ver a equivalência aproximada entre as duas quantidades, apenas divergem para $k \approx k_{cutoff}$. Que corresponde ao grau dos primeiros nodos que não respeitam a condição $1 \ll \nu$. Estas redes têm populações $N = 50\,000$.

Tomaram-se três tamanhos de rede diferentes, $N = \{1; 1,5; 2\} \times 10^5$, e cinco atracções adicionais, $A = \{-0,5; -0,2; 0; 0,5; 1,5\}$. Após o tratamento de dados, segundo os métodos apresentados anteriormente nesta secção, medem-se os valores dos expoentes γ , η e δ . Nas Figuras 8, 9 e 10 mostram-se os gráficos de $F(k)$, $F(g)$ e $G(k)$ respectivamente, para redes com populações $N = \{1,5; 2\} \times 10^5$ e atracção adicional $A = 0$, o que equivale ao modelo B-A.

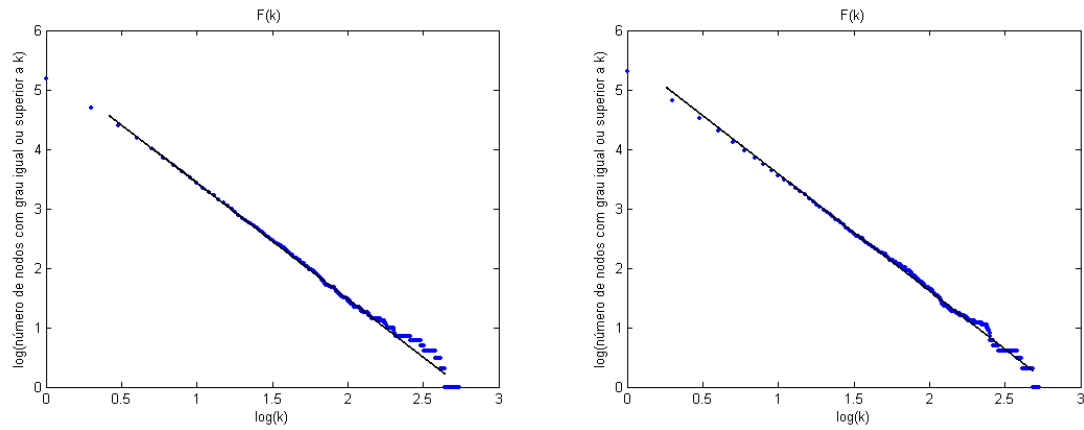


Figura 8: Logaritmo do número de vértices com grau superior a k em função do logaritmo de k . O gráfico da esquerda é de uma rede com $N = 1,5 \times 10^5$, a linha preta tem um declive de $\approx -1,96$ o que traduz um $\gamma \approx 2,96$. O da direita é para $N = 2 \times 10^5$, e tem $\gamma \approx 2,95$.

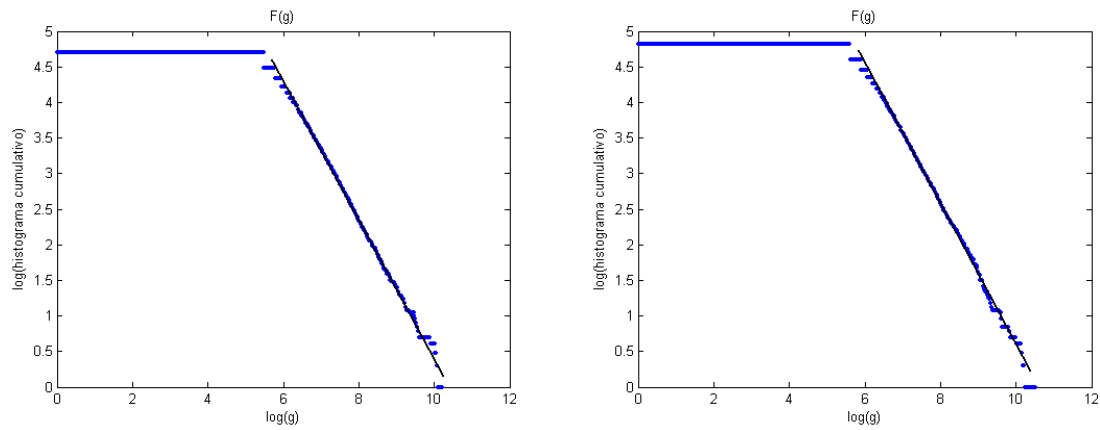


Figura 9: Histograma Cumulativo da distribuição de *betweenness*. O gráfico da esquerda é de uma rede com $N = 1,5 \times 10^5$, a linha preta tem um declive de $\approx -0,98$ o que corresponde a um $\delta \approx 1,98$. O da direita é para $N = 2 \times 10^5$, e apresenta $\delta \approx 1,99$.

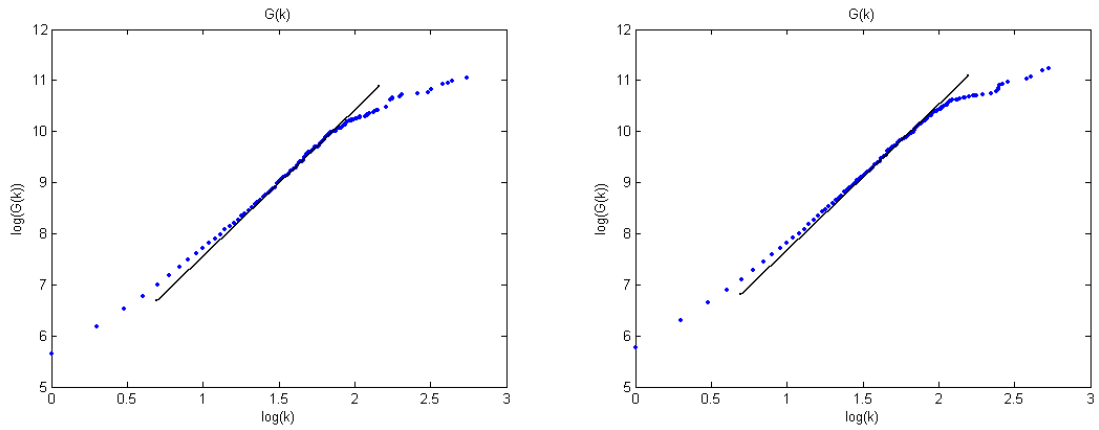


Figura 10: Função $G(k)$. O gráfico da esquerda é de uma rede com $N = 1,5 \times 10^5$, a linha preta tem um declive de $\approx 2,94$ o que representa um $\eta \approx 1,94$. O da direita é para $N = 2 \times 10^5$, e exibe um $\eta \approx 1,96$.

Os resultados para os restantes valores tomados pela constante de atracção adicional são visíveis nas linhas das Tabelas I, II e III. O número de realizações R para cada valor de N e A , é indicado numa coluna das próprias tabelas. A média dos expoentes γ e η sobre as várias realizações é apresentada numa coluna separada com nomes $\bar{\gamma}$ e $\bar{\eta}$.

Tabela I:

 $N = 10^5$

A	R	γ	$\bar{\gamma}$	η	$\bar{\eta}$	δ	$\gamma - 1 - \eta$
-0,5	5	2,48	2,51	1,37	1,40	1,99	0,11
		2,46*		1,49*		1,96*	-0,031
		2,51		1,37		2,00	0,15
		2,46		1,43		1,96	0,023
		2,52		1,33		2,01	0,20
-0,2	5	2,73	2,81	1,66	1,70	1,98	0,074
		2,81		1,77		1,98	0,040
		2,87		1,67		1,99	0,18
		2,81		1,70		1,98	0,11
		2,81		1,71		2,01	0,097
0	5	2,93	2,97	1,84	1,82	1,97	0,094
		3,00*		1,82*		2,00*	0,18
		2,98		1,94		1,97	0,042
		2,99		1,76		2,02	0,22
		2,94		1,72		1,99	0,22
0,5	5	3,43	3,38	1,99	2,29	1,99	0,44
		3,38		2,11		2,02	0,26
		3,38		1,97		1,98	0,41
		3,48		2,26		2,01	0,24
		3,21		2,14		1,96	0,061
1,5	5	4,30	4,02	2,44	2,59	1,95	0,85
		3,89		2,50		2,00	0,39
		3,99*		2,67*		1,98*	0,32
		3,89		2,78		1,96	0,051
		4,01		2,58		2,02	0,42

Tabela II:

$N = 1,5 \times 10^5$

A	R	γ	$\bar{\gamma}$	η	$\bar{\eta}$	δ	$\gamma - 1 - \eta$
-0,5	5	2,50*	2,49	1,47*	1,46	1,98*	0,027
		2,44		1,42		1,97	0,018
		2,54		1,32		1,98	0,21
		2,51		1,35		2,00	0,16
		2,48		1,46		1,98	0,021
-0,2	5	2,78	2,77	1,64	1,64	1,98	0,13
		2,71		1,54		1,98	0,17
		2,81		1,82		1,97	-0,012
		2,74		1,55		2,02	0,19
		2,82		1,66		1,98	0,16
0	5	2,98	2,95	2,01	1,81	2,01	-0,036
		2,95		1,77		1,99	0,18
		2,96		1,65		1,99	0,31
		2,96*		1,84*		1,98*	0,12
		2,88		1,80		1,99	0,086
0,5	5	3,29	3,39	2,24	2,19	2,00	0,049
		3,46		2,13		2,02	0,32
		3,38		2,25		1,98	0,13
		3,42		2,17		1,98	0,25
		3,40		2,17		1,97	0,23
1,5	5	3,95	4,05	2,49	2,71	2,00	0,46
		4,24		2,65		2,00	0,56
		4,14		2,93		1,99	0,21
		4,07*		2,71*		1,99*	0,36
		3,85		2,78		2,00	0,066

Tabela III:

$$N = 2 \times 10^5$$

A	R	γ	$\bar{\gamma}$	η	$\bar{\eta}$	δ	$\gamma - 1 - \eta$
-0,5	2	2,51*	2,50	1,46*	1,47	2,02*	0,048
		2,47		1,44		2,00	0,032
		2,52		1,51		1,98	0,0017
-0,2	2	2,76	2,79	1,76	1,72	1,99	0,017
		2,79		1,74		1,97	0,045
		2,83		1,66		2,02	0,18
0	2	2,95*	2,97	1,86*	1,86	1,99*	0,10
		2,95		1,86		1,99	0,088
		3,01		1,86		2,03	0,15
0,5	2	3,36	3,44	2,32	2,29	2,00	0,031
		3,48		2,25		1,97	0,23
		3,47		2,30		1,98	0,17
1,5	2	4,13*	4,05	3,17*	2,93	1,99*	0,040
		3,95		2,84		2,00	0,11
		4,06		2,78		1,98	0,28

Tamanho das redes que se toma para as simulações é limitado pelo tempo necessário ao cálculo da BC, mesmo com o algoritmo cumulativo proposto em [16] e apresentado anteriormente o tempo necessário ao cálculo da BC é proporcional a N^2 o que torna o processo extremamente moroso para N 's elevados. Por exemplo, o tempo necessário para efectuar o cálculo em cada uma das maiores estruturas aqui analisadas, com $N = 2 \times 10^5$, é de cerca de duas semanas. Esta característica do algoritmo proíbe cálculos para tamanhos razoavelmente superiores, uma estimativa do tempo necessário ao cálculo da BC numa árvore com $N = 1 \times 10^6$, ou seja cinco vezes maior, é cerca de um ano. Para uma estrutura com tamanho uma ordem de grandeza maior é cerca de quatro anos.

* aos valores assinalados correspondem os gráficos de dados típicos desenhados em Anexo

4-DISCUSSÃO

O estudo apresentado neste documento aborda o tema da centralidade em árvores evolutivas *scale-free* recorrendo à técnica da aproximação ao contínuo [1,8]. O emprego deste método permite calcular a quantidade de vértices em cada uma das regiões que são ligadas somente por v . Mostra-se que, nas condições de $1 \ll v \ll t$, todas estas regiões crescem com t/v à excepção de uma, a dos ascendentes, essa cresce linearmente com o tempo.

Por este motivo a definição de BC (39) é simplificada, e facilmente se verifica que os cálculos conduzem à relação $g \sim t^2 v^{-1} \sim t k^{\gamma-1}$. Isto fornece uma sustentação teórica para as observações numéricas feitas em [12,15] da relação $g \sim k^\eta$, e acrescenta a relação $\eta = \gamma - 1$. Levantando-se assim a dependência do expoente η com os detalhes da estrutura das árvores *scale-free*.

Este resultado reflecte a independência da distribuição de BC em relação à distribuição de grau. No limite de $\gamma \rightarrow 2$ a relação entre centralidade e grau é linear. Quanto maior for γ mais forte é a dependência da BC com a conectividade, porém menor é a variância da conectividade, pois $P(k) \sim k^{-\gamma}$. Na realidade, mostra-se que os dois efeitos se anulam, ou seja, a distribuição de BC é insensível à distribuição de grau, e é $P(g) \sim g^{-2}$. Isto está em consonância com os resultados numéricos e analítico encontrados em [13]. A razão deste comportamento é que para a distribuição de BC as regras de preferência são irrelevantes, a BC de um vértice é determinada pelo momento em que nasceu. Esta propriedade topológica decorre do facto de os vértices mais antigos ocuparem os lugares mais centrais da estrutura, porque independentemente do seu grau ligam vários grandes componentes. Por outro lado, os mais recentes ainda que tenham grau equiparável ao dos mais antigos ligam apenas um grande componente a componentes pequenos, a maior parte dos caminhos acaba por ficar dentro do grande componente.

É possível depreender a relação $\eta = \gamma - 1$ com base na noção de insensibilidade da distribuição de BC aos detalhes da estrutura. Olhando para a dedução feita em 1.2.3 vê-se que se a distribuição de BC não sente os detalhes da estrutura, ou seja, $\delta = 2$

independentemente do expoente γ , a relação (24) sofre uma importante simplificação e aparece reduzida a $g \sim k^{\gamma-1}$.

É sugerido em [13], com base em resultados numéricos, que as distribuições de BC dos vértices g_n e das arestas g_v têm comportamentos semelhantes. Mostra-se analiticamente, na secção 2.4, que $g_n(v)$ é dominada por $g_v(e_1, v)$ de tal forma que $g_n(v) \approx g_v(e_1, v)$, para $1 \ll v \ll t$. Recordando que, no modelo estudado, em cada passo de tempo se junta à rede um novo par vértice e aresta, vemos que a relação é recíproca. Isto é, a BC de cada aresta é aproximadamente igual à BC do nodo que nasceu na mesma iteração, e vice-versa. Posto isto, é imediato que as suas distribuições são aproximadamente iguais, $P(g_n) \approx P(g_v)$.

As deduções aqui apresentadas são exactas apenas para $N \rightarrow \infty$, o que leva os resultados numéricos e serem tanto mais próximos dos previstos quanto maior o tamanho das redes simuladas. Pode observar-se o comportamento manifestamente *scale-free* da distribuição de grau cumulativa $F(k)$ das estruturas simuladas nos gráficos desenhados em Anexo.

A observação das tabelas de resultados confirma fortemente que a previsão $\delta = 2$ para todo $\gamma \in]2, \infty[$, para os três N 's considerados. Pode ver-se que os valores de δ estimados pela estatística utilizada são muito próximos do previsto, com desvios da ordem de 1%. Gráficos representando as distribuições de BC cumulativas $F(g)$ típicas das redes simuladas são apresentados em Anexo. Pode ver-se nos gráficos do Anexo que $F(g)$ é idêntica para as redes com o mesmo N , quase independentemente da atracção adicional A . A diferenças no limite superior do eixo vertical nos gráficos $F(g)$ devem-se aos diferentes números de vértices com centralidade nula ($k = 1$) para diferentes valores de A . O comportamento de δ aparenta ser efectivamente como o previsto, e parece manifestar-se logo a partir tamanhos menores do que os aqui considerados.

A verificação numérica da relação entre a BC e a conectividade $g \sim k^{\gamma-1}$ é um pouco menos robusta. Podemos ver nos gráficos do Anexo que o comportamento de $G(k)$ se desvia um pouco do previsto para os N 's menores, principalmente fora dos limites de

$1 \ll k \ll k_{cutoff}$, o que, tendo em conta a dinâmica da rede corresponde aos limites $1 \ll v \ll t$. Assim apenas os valores da zona $1 \ll k \ll k_{cutoff}$ são considerados na linearização. Porém para $N = 2 \times 10^5$ os desvios são já bastante pequenos denotando uma boa adequação ao modelo teórico desenvolvido neste trabalho. Os gráficos típicos de $G(k)$ são apresentados também em Anexo.

Podemos observar na coluna mais à direita das Tabelas I, II e III que o desvio relativo do expoente η da previsão aqui feita aumenta à medida que γ aumenta. O que pode dever-se ao facto de, apesar de BC ter sensivelmente a mesma distribuição, o k_{cutoff} decresce ordens de grandeza entre $A = -0.5$ e $A = 1.5$ para os N 's considerados, como se pode ver nos gráficos em Anexo, intensificando os efeitos de tamanho finito.

Nota-se ainda, por comparação entre Tabelas, uma confluência para os valores previstos à medida que os tamanhos tomados para as redes simuladas aumentam. O que é um indicio de que para estruturas grandes o suficiente os desvios se tornariam desprezáveis. No entanto, a execução de um maior número de realizações poderia contribuir para a comprovação da validade da abordagem aqui proposta.

5-CONCLUSÕES

A teoria descrita neste trabalho sugere que o expoente que relaciona a BC e a conectividade dos vértices das árvores *scale-free*, η , depende do expoente *scale-free* γ de acordo com a relação $\eta = \gamma - 1$. Esta afirmação é suportada pelas simulações executadas, contudo uma expansão do tamanho das redes simuladas e do número de realizações para cada constante de atracção adicional pode produzir provas menos discutíveis.

Conclui-se inequivocamente, via analítica e numérica, que a distribuição de BC neste tipo de redes é independente de todos os parâmetros da estrutura, tomando invariavelmente a forma de uma lei de potência com expoente $\delta = 2$. Mostra-se ainda como a BC das arestas tratada em [13], e a dos vértices tratada neste trabalho (e convencional) são aproximadamente equivalentes.

6-ANEXO

Nesta secção são apresentados gráficos de dados característicos obtidos das simulações executadas, a partir dos quais se calcularam os resultados exibidos nas tabelas. Em cada figura encontram-se desenhadas representações logarítmicas das funções cumulativas (55), (56) e (57) referidas na secção 3, nomeadamente $F(k)$, $F(g)$ e $G(k)$, e as suas respectivas rectas de ajuste. Mostram-se os dados para uma realização de cada combinação das populações $N = \{1; 1,5; 2\} \times 10^5$ com as atracções adicionais, $A = \{-0,5; 0; 1,5\}$.

Os eixos verticais não se encontram normalizados, o que não tem qualquer influência nos valores calculados para os expoentes. Pois o expoente da lei de potência é indicado directamente pelo declive da recta de ajuste ao gráfico logarítmico usando as definições (55), (56) e (57), e qualquer constante de normalização somente influenciaria a ordenada na origem da recta de ajuste.

Em cada figura o gráfico de $F(k)$, assinalado com a), representa o logaritmo do número de vértices com grau igual ou superior a k em função do logaritmo de k . Os gráficos $F(g)$, assinalados com b), são a representação o logarítmica da função cumulativa do histograma do número de vértices que têm BC num certo intervalo, em função do logaritmo do valor de BC no centro desse intervalo. Os gráficos assinalados com c), de $G(k)$, representam a função cumulativa do logaritmo da centralidade média dos vértices com conectividade k em função do logaritmo de k .

Os valores calculados, através das rectas de ajuste, para γ , δ e η nas figuras aqui apresentadas encontram-se assinalados com * nas Tabelas I, II e III da secção 3.

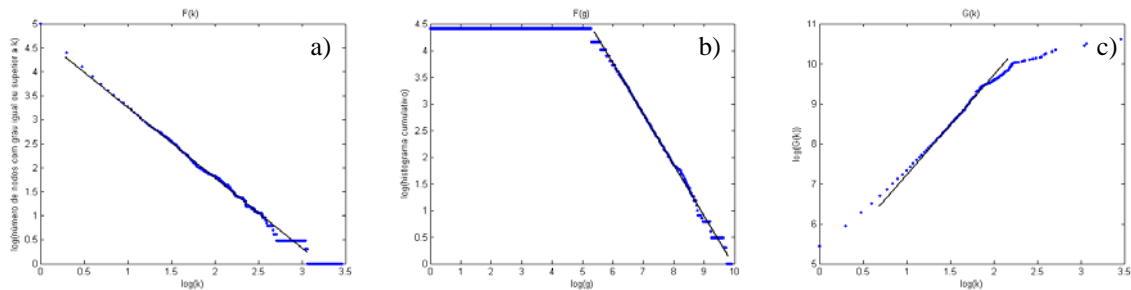


Figura A1: Dados oriundos da simulação de uma rede de tamanho $N = 1 \times 10^5$ e constante de atracção adicional $A = -0.5$.

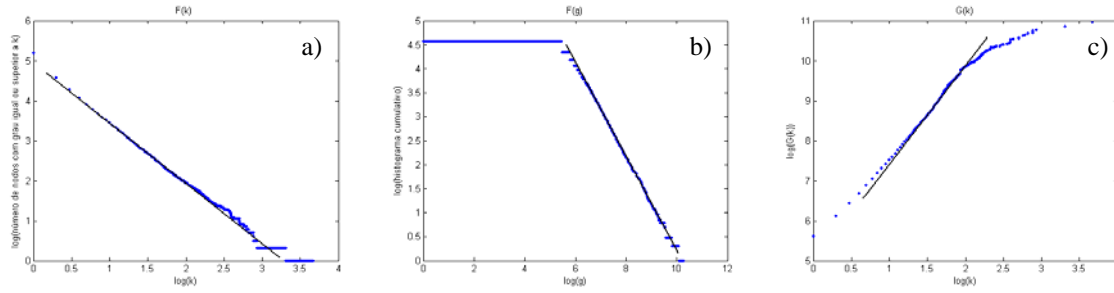


Figura A2: Dados da simulação de uma rede de tamanho $N = 1,5 \times 10^5$ e constante de atracção adicional $A = -0.5$.

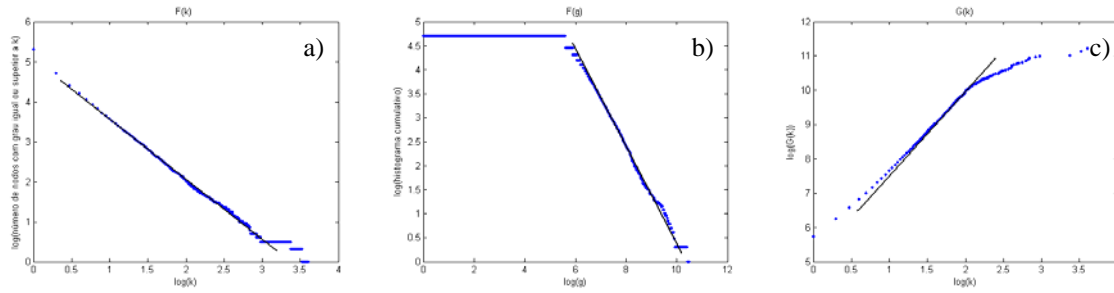


Figura A3: Dados da simulação de uma rede de tamanho $N = 2 \times 10^5$ e constante de atracção adicional $A = -0.5$.

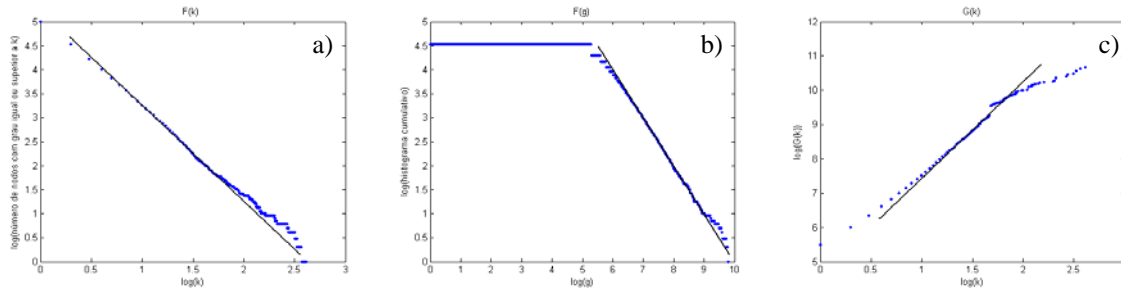


Figura A4: Dados da simulação de uma rede de tamanho $N = 1 \times 10^5$ e constante de atracção adicional $A = 0$.

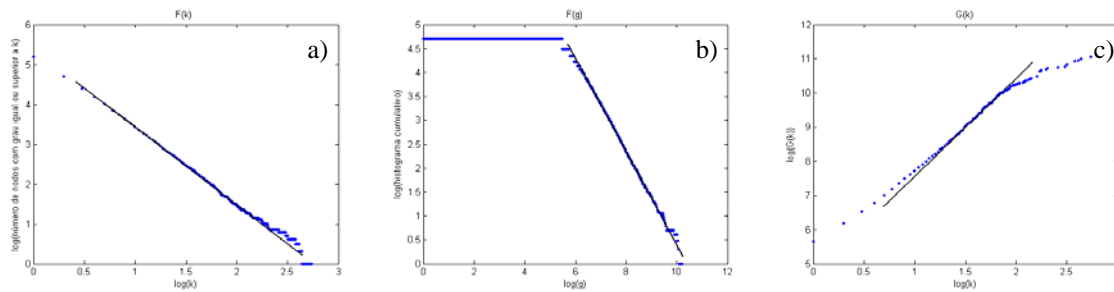


Figura A5: Dados da simulação de uma rede de tamanho $N = 1,5 \times 10^5$ e constante de atracção adicional $A = 0$.

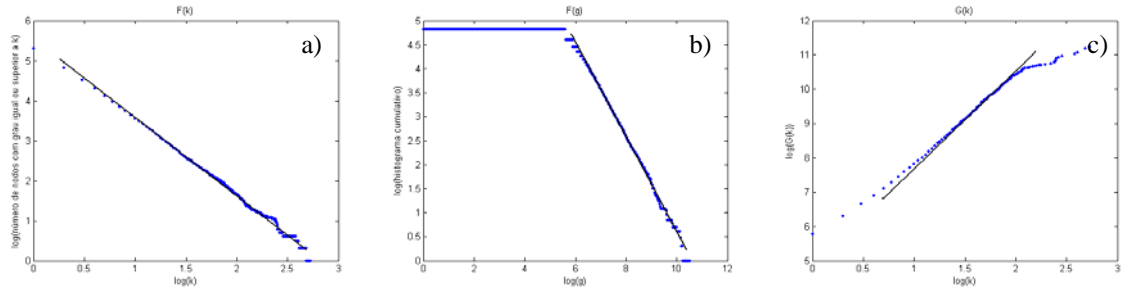


Figura A6: Dados da simulação de uma rede de tamanho $N = 2 \times 10^5$ e constante de atracção adicional $A = 0$.

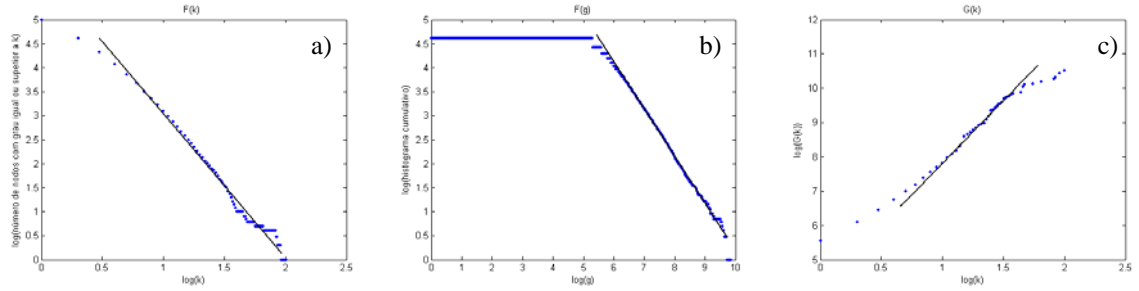


Figura A7: Dados da simulação de uma rede de tamanho $N = 1 \times 10^5$ e constante de atracção adicional $A = 1.5$.

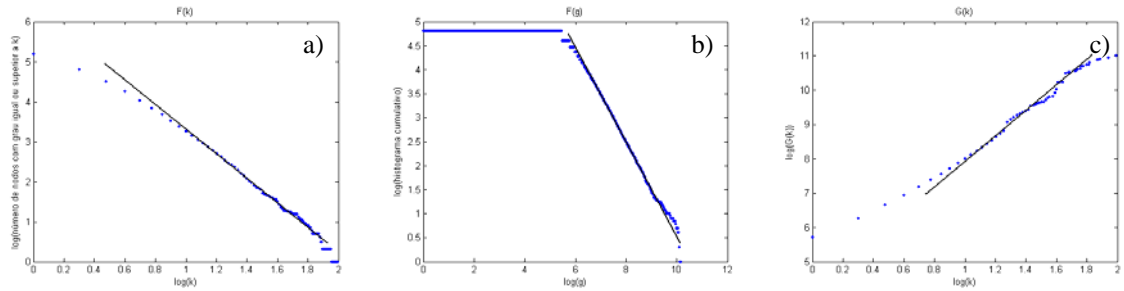


Figura A8: Dados da simulação de uma rede de tamanho $N = 1,5 \times 10^5$ e constante de atracção adicional $A = 1.5$.

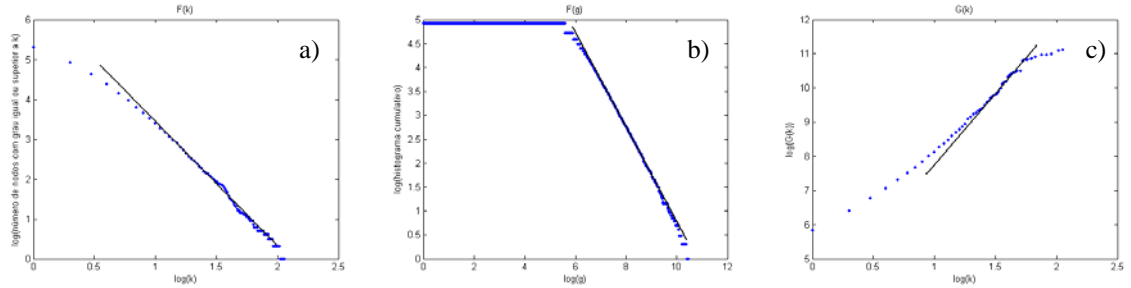


Figura A9: Dados da simulação de uma rede de tamanho $N = 2 \times 10^5$ e constante de atracção adicional $A = 1.5$.

7-REFERÊNCIAS

- [1] Dorogovtsev, S.N., Mendes, J.F.F.,
Evolution of Networks,
Adv. Phys. 51, 1079 (2002)
[cond-mat/0106144](#)
- [2] Dorogovtsev S. N., Mendes, J. F. F. Samukhin, A. N.,
Anomalous Percolating Properties of Growing Networks,
Phys. Rev. E 64, 066110 (2001)
[cond-mat/0106141](#)
- [3] Barabási, A.-L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., Vicsek T.,
Evolution of the Social Network of Scientific Collaborations,
Physica A 311, (3-4) (2002)
[cond-mat/0104162](#)
- [4] Barabási, A.-L. and Albert, R.,
Emergence of Scaling in Random Networks,
Science 286, 509 (1999)
[cond-mat/9910332](#)
- [5] Barabási, A.-L., Albert, R., Jeong, H., ,
Scale-Free Characteristics of Random Networks:The Topology of the World
Wide Web,
Physica A, 281, 69 (2000)
- [6] Dorogovtsev, S. N., Mendes, J. F. F., Samukhin, A. N., 2000,
Structure of Growing Networks With Preferential Linking,
Phys. Rev. Lett., 85, 4633 (2000)
[cond-mat/0004434](#)
- [7] Dorogovtsev, S. N., Mendes, J. F. F.,
Evolution of Networks With Aging of Sites,
Phys. Rev. E, 62, 1842 (2000)
[cond-mat/0001419](#)
- [8] Dorogovtsev, S. N., Mendes, J. F. F.,
Scaling Properties of Scale-Free Evolving Networks: Continuous Approach,
Phys. Rev. E, 63, 056125 (2001)
[cond-mat/0012009](#)
- [9] Albert, R. and Barabási, A.-L.,
Statistical Mechanics of Complex Networks,
Rev. Mod. Phys. 74, 47 (2002)
[cond-mat/0106096](#)

- [10] M.E.J. Newman,
Scientific Collaboration Networks,
Phys. Rev. E 64, 016131 (2001); Phys. Rev. E 64, 016132 (2001)
[cond-mat/0011144](#)
- [11] L. C. Freeman,
A Set of Measures of Centrality Based on Betweenness,
Sociometry 40, 35 (1977)
- [12] K.-I. Goh, B. Kahng, and D. Kim,
Universal Behavior of Load Distribution in Scale-free Networks,
Phys. Rev. Lett. 87, 278701 (2001)
[cond-mat/0106565](#)
- [13] Goh, K.-I., Jeong, H., Kahng, B., Kim, D.,
Classification of Scale-Free Networks,
Proc. Natl. Acad. Sci. USA 99, 12583 (2002)
[cond-mat/0205232](#)
- [14] M. Barthélemy,
Betweenness Centrality in Large Complex Networks,
Eur. Phys. Jour. B, vol 38, 163 (2004)
[cond-mat/0309436](#)
- [15] Vazquez, A., Pastor-Satorras, R., Vespignani, A.,
Large-scale topological and dynamical properties of Internet
Phys. Rev. E 65, 066130 (2002)
[cond-mat/0112400](#)
- [16] U. Brandes,
A Faster Algorithm for Betweenness Centrality,
Journal of Mathematical Sociology, 25, 163 (2001)
[inf.uni-konstanz.de/algo/publications/b-fabc-01.pdf](#)
- [17] Krapivsky, P. L., Redner, S., Leyvraz, F.,
Connectivity of Growing Random Networks,
Phys. Rev. Lett. 85, 4629-4632 (2000)
[cond-mat/0005139](#)